

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: COSGROVE, DANIEL J.;

GUILTINAN, MARK;

5 SHCHERBAN, TATYANA;

SHI, JUN

(ii) TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

10 (A) INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
STATE UNIVERSITY

(B) STREET: 113 TECHNOLOGY CENTER

(C) CITY: UNIVERSITY PARK

(D) STATE: PENNSYLVANIA

15 (E) COUNTRY: UNITED STATES OF AMERICA

(F) ZIP: 16802-7000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: FLOPPY DISK

(B) COMPUTER: NEC 286

20 (C) OPERATING SYSTEM: DOS

(D) SOFTWARE: WORDPERFECT 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

25 (C) CLASSIFICATION:

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 681

(B) TYPE: NUCLEIC ACID

30 (C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: UNKNOWN

Seq # 4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

35 GAC TAC GGT GGC TGG CAG AGC GGC CAC GCC ACC TTT TAT GGT 42
Asp Tyr Gly Gly Trp Gln Ser Gly His Ala Thr Phe Tyr Gly
1 5 10

GGT GGT GAC GCA TCT GGC ACC ATG GGT GGA GCT TGT GGG TAT 84
Gly Gly Asp Ala Ser Gly Thr Met Gly Gly Ala Cys Gly Tyr
15 20 25

40 GGG AAT TTA TAC AGC CAA GGG TAT GGC ACG AAC ACG GTG GCG 126
Gly Asn Leu Tyr Ser Gln Gly Tyr Gly Thr Asn Thr Val Ala
30 35 40

CTG AGC ACT GCG CTA TTT AAC AAT GGA TTA AGT TGT GGT GCT 168
Leu Ser Thr Ala Leu Phe Asn Asn Gly Leu Ser Cys Gly Ala
45 50 55

45 TGC TTC GAA ATG ACT TGT ACA AAC GAC CCT AAA TGG TGC CTT 210
Cys Phe Glu Met Thr Cys Thr Asn Asp Pro Lys Trp Cys Leu
60 65 70

CCG GGA ACT ATT AGG GTC ACT GCC ACC AAC TTT TGC CCT CCT 252
Pro Gly Thr Ile Arg Val Thr Ala Thr Asn Phe Cys Pro Pro
50 75 80

AAC TTT GCT CTC CCT AAC AAC AAT GGT GGA TGG TGC AAC CCT 294
Asn Phe Ala Leu Pro Asn Asp Asp Gly Gly Trp Cys Asn Pro
85 90 95

55 CCT CTC CAA CAC TTC GAC ATG GCT GAG CCT GCC TTC CTT CAA 336
Pro Leu Gln His Phe Asp Met Ala Glu Pro Ala Phe Leu Gln
100 105 110

ATC GCT CAA TAC CGA GCT GGT ATC GTC CCC GTC TCC TTT CGT 378
Ile Ala Gln Tyr Arg Ala Gly Ile Val Pro Val Ser Phe Arg
115 120 125

60 AGG GTA CCA TGT ATG AAG AAA GGT GGA GTG AGG TTT ACA ATC 420
Arg Val Pro Cys Met Lys Lys Gly Gly Val Arg Phe Thr Ile
130 135 140

65 AAT GGC CAC TCA TAC TTC AAC CTC GTT TTG ATC ACA AAC GTC 462
Asn Gly His Ser Tyr Phe Asn Leu Val Leu Ile Thr Asn Val
145 150

GGT GGC GCA GGC GAC GTC CAC TCT GTG TCG ATA AAG GGG TCT 504
Gly Gly Ala Gly Asp Val His Ser Val Ser Ile Lys Gly Ser
155 160 165

70 CGA ACT GGA TGG CAA TCC ATG TCT AGA AAT TGG GGC CAA AAC 546
Arg Thr Gly Trp Gln Ser Met Ser Arg Asn Trp Gly Gln Asn
170 175 180

TGG CAA AGC AAC AAC TAT CTC AAT GGC CAA GGC CTT TCC TTT 588
Trp Gln Ser Asn Asn Tyr Leu Asn Gly Gln Gly Leu Ser Phe
185 190 195

75 CAA GTC ACT CTT AGT GAT GGT CGC ACT CTC ACT GCC TAT AAT 630
Gln Val Thr Leu Ser Asp Gly Arg Thr Leu Thr Ala Tyr Asn
200 205 210

80 CTC GTT CCT TCC AAT TGG CAA TTT GGC CAA ACC TAT GAA GGC 672
Leu Val Pro Ser Asn Trp Gln Phe Gly Gln Thr Tyr Glu Gly
215 220

CCT CAA TTC 681
Pro Gln Phe
225

85 (3) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228
(B) TYPE: AMINO ACID
(D) TOPOLOGY: UNKNOWN
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

90 Ala Gly Gly Gly Trp Val Asn Ala His Ala Thr Phe Tyr Gly Gly
1 5 10 15

Gly Asp Ala Ser Gly Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn
20 25 30

95 Leu Tyr Ser Gln Gly Tyr Gly Thr Asn Thr Ala Ala Leu Ser Thr
35 40 45

Ala Leu Phe Asn Asn Gly Leu Ser Cys Gly Ala Cys Phe Glu Ile
50 55 60

	Arg Cys Gln Asn Asp Gly Lys Trp Cys Leu Pro Gly Ser Ile Val	
	65	70 75
100	Val Thr Ala Thr Asn Phe Cys Pro Pro Asn Asn Ala Leu Pro Asn	
	80	85 90
	Asn Ala Gly Gly Trp Cys Asn Pro Pro Gln Gln His Phe Asp Leu	
	95	100 105
105	Ser Gln Pro Val Phe Gln Arg Ile Ala Gln Tyr Arg Ala Gly Ile	
	110	115 120
	Val Pro Val Ala Tyr Arg Arg Val Pro Cys Val Arg Arg Gly Gly	
	125	130 135
	Ile Arg Phe Thr Ile Asn Gly His Ser Tyr Phe Asn Leu Val Leu	
	140	145 150
110	Ile Thr Asn Val Gly Gly Ala Gly Asp Val His Ser Ala Met Val	
	155	160 165
	Lys Gly Ser Arg Thr Gly Trp Gln Ala Met Ser Arg Asn Trp Gly	
	170	175 180
115	Gln Asn Trp Gln Ser Asn Ser Tyr Leu Asn Gly Gln Ser Leu Ser	
	185	190 195
	Phe Lys Val Thr Thr Ser Asp Gly Gln Thr Ile Val Ser Asn Asn	
	200	205 210
	Xaa Ala Asn Ala Gly Trp Ser Phe Gly Gln Thr Phe Thr Gly Ala	
	215	220 225
120	His Val Arg	

(4) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 222

(B) TYPE: AMINO ACID

(D) TOPOLOGY: UNKNOWN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

His Met Gly Pro Trp Ile Asn Ala His Ala Thr Phe Tyr Xaa Xaa
1 5 10 15

130 Gly Asp Ala Xaa Xaa Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn
20 25 30

Leu Tyr Ser Gln Gly Tyr Gly Leu Glu Thr Ala Ala Leu Ser Thr
35 40 45

Ala Leu Phe Asp Gln Gly Leu Ser Cys Gly Ala Cys Xaa Glu Leu
50 55 60

135 Met Cys Val Asn Asp Pro Gln Trp Cys Ile Lys Gly Arg Ser Ile
65 70 75

Val Val Thr Ala Thr Asn Phe Cys Pro Pro Gly Gly Ala Cys Asp
80 85 90

140 Pro Pro Asn His His Phe Asp Leu Ser Gln Pro Ile Tyr Glu Lys
95 100 105

Ile Ala Leu Tyr Lys Ser Gly Ile Ile Pro Val Met Tyr Arg Arg
110 115 120

Val Arg Cys Lys Arg Ser Gly Gly Ile Arg Phe Thr Ile Asn Gly
125 130 135

145 His Ser Tyr Phe Asn Leu Val Leu Val Thr Asn Val Gly Gly Ala
140 145 150

Gly Asp Val His Ser Val Ser Met Lys Gly Ser Arg Thr Lys Trp
155 160 165

150 Gln Leu Met Ser Arg Asn Trp Gly Gln Asn Trp Gln Ser Asn Ser
170 175 180

Tyr Leu Asn Gly Gln Ser Leu Ser Phe Val Val Thr Thr Ser Asp
185 190 195

Arg Arg Ser Val Val Ser Phe Asn Val Ala Pro Pro Thr Trp Ser
200 205 210

155 Phe Gly Gln Thr Tyr Thr Gly Gly Gln Phe Arg Tyr
215 220

(5) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

160 (A) LENGTH: 227
 (B) TYPE: AMINO ACID
 (D) TOPOLOGY: UNKNOWN
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

165 Lys XAA Ser Val Ala Gln Ser Ala Phe Ala Thr Phe Tyr Gly Gly
 1 5 10 15

Lys Asp Gly Ser Cys Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn
 20 25 30

170 Leu Tyr Asn Ala Gly Tyr Gly Leu Tyr Asn Ala Ala Leu Ser Ser
 35 40 45

Ala Leu Phe Asn Asp Gly Ala Met Cys Gly Ala Cys Tyr Thr Ile
 50 55 60

Thr Cys Asp Thr Ser Gln Thr Lys Trp Cys Lys Pro Gly Gly Asn
 65 70 75

175 Ser Ile Thr Ile Thr Ala Thr Asn Leu Cys Xaa Pro Asn Trp Ala
 80 85 90

Leu Pro Ser Asn Ser Gly Gly Trp Cys Asn Pro Pro Leu Xaa His
 95 100 105

180 Phe Asp Met Ser Gln Pro Ala Trp Glu Asn Ile Ala Val Tyr Gln
 110 115 120

Ala Gly Ile Val Pro Val Asn Tyr Lys Arg Val Pro Xaa Gln Arg
 125 130 135

Ser Gly Gly Ile Arg Phe Ala Ile Ser Gly His Asp Tyr Phe Glu
 140 145 150

185 Leu Val Thr Val Thr Asn Val Gly Gly Ser Gly Val Val Ala Gln
 155 160 165

Met Ser Ile Lys Gly Ser Asn Thr Gly Trp Met Ala Met Ser Arg
 170 175 180

190 Asn Trp Gly Ala Asn Trp Gln Ser Asn Ala Tyr Leu Ala Gly Gln
 185 190 195

Ser	Leu	Ser	Phe	Ile	Val	Gln	Leu	Asp	Asp	Gly	Arg	Lys	Val	Thr
				200					205					210

Ala	Trp	Asn	Xaa	Ala	Pro	Xaa	Asn	Trp	Leu	Xaa	Xaa	Xaa	Xaa	Xaa
				215					220					225

195 Xaa Xaa

(6) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 225

(B) TYPE: AMINO ACID

200 (D) TOPOLOGY: UNKNOWN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Asp	Asn	Gly	Gly	Trp	Glu	Arg	Gly	His	Ala	Thr	Phe	Tyr	Gly	Gly
1				5					10					15

Ala	Asp	Ala	Ser	Gly	Thr	Met	Gly	Gly	Ala	Cys	Gly	Tyr	Gly	Asn
				20					25					30

Leu	His	Ser	Gln	Gly	Tyr	Gly	Leu	Gln	Thr	Ala	Ala	Leu	Ser	Thr
				35					40					45

Ala	Leu	Phe	Asn	Ser	Gly	Gln	Lys	Cys	Gly	Ala	Cys	Phe	Glu	Leu
				50					55					60

Thr	Cys	Glu	Asp	Asp	Pro	Glu	Trp	Cys	Ile	Pro	Gly	Ser	Ile	Ile
				65					70					75

Val	Arg	Tyr	Asn	Leu	Ala	Asn	Phe	Ala	Leu	Ala	Asn	Asp	Asn	Gly
				80					85					90

Gly	Trp	Cys	Asn	Pro	Pro	Leu	Lys	His	Phe	Asp	Leu	Ala	Glu	Pro
				95					100					105

Ala	Phe	Leu	Gln	Ile	Ala	Gln	Tyr	Arg	Ala	Gly	Ile	Val	Pro	Val
				110					115					120

Ala	Phe	Arg	Arg	Val	Pro	Cys	Glu	Lys	Gly	Gly	Gly	Ile	Arg	Phe
				125					130					135

Thr	Ile	Asn	Gly	Asn	Pro	Tyr	Phe	Asp	Leu	Val	Leu	Ile	Thr	Asn
				140					145					150

Val	Gly	Gly	Ala	Gly	Asp	Ile	Arg	Ala	Val	Ser	Leu	Lys	Gly	Ser
				155					160					165

225 Lys Thr Asp Gln Trp Gln Ser Met Ser Arg Asn Trp Gly Gln Asn
170 175 180

Trp Gln Ser Asn Thr Tyr Leu Arg Gly Gln Ser Leu Ser Phe Gln
185 190 195

Val Thr Asp Ser Asp Gly Arg Thr Val Val Ser Tyr Asp Val Val
200 205 210

230 Pro His Asp Trp Gln Phe Gly Gln Thr Phe Glu Gly Gly Gln Phe
215 220 225

(7) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

235 (A) LENGTH: 226

(B) TYPE: AMINO ACID

(D) TOPOLOGY: UNKNOWN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

240 Asp Tyr Ser Ser Trp Gln Ser Ala His Ala Thr Phe Tyr Gly Gly
1 5 10 15

Gly Asp Ala Ser Gly Thr Met Gly Gly Thr Cys Gly Tyr Gly Asn
20 25 30

Leu Tyr Ser Thr Gly Tyr Thr Asn Thr Ala Ala Leu Ser Thr Val
35 40 45

245 Leu Phe Asn Asp Gly Ala Ala Cys Arg Ser Cys Tyr Glu Leu Arg
50 55 60

Cys Asp Asn Asp Gly Gln Trp Cys Leu Pro Gly Ser Val Thr Val
65 70 75

250 Thr Ala Thr Asn Leu Cys Pro Pro Asn Tyr Ala Leu Pro Asn Asp
80 85 90

Asp Gly Gly Trp Cys Asn Pro Pro Arg Pro His Phe Asp Met Ala
95 100 105

Glu Pro Ala Phe Leu Gln Ile Gly Val Tyr Arg Ala Gly Ile Val
110 115 120

255 Pro Val Ser Tyr Arg Arg Val Pro Cys Val Lys Lys Gly Gly Ile
125 130 135

Arg Phe Thr Ile Asn Gly His Ser Tyr Phe Asn Leu Val Leu Val
140 145 150

260 Thr Asn Val Ala Gly Pro Gly Asp Val Gln Ser Val Ser Ile Lys
155 160 165

Gly Ser Ser Thr Gly Trp Gln Pro Met Ser Arg Asn Trp Gly Gln
170 175 180

Asn Trp Gln Ser Asn Ser Tyr Leu Asp Gly Gln Ser Leu Ser Phe
185 190 195

265 Gln Val Ala Val Ser Asp Gly Arg Thr Val Thr Ser Asn Asn Val
200 205 210

Val Pro Ala Gly Trp Gln Phe Gly Gln Thr Phe Glu Gly Gly Gln
215 220 225

Phe